Mathematical Statistician and Engineering Applications
ISSN: 2094-0343
DOI: https://doi.org/10.17762/msea.v70i2.2328

Enhancing MNIST Digit Recognition with Ensemble Learning Techniques

Divya Kapil

Asst. Professor, School of Computing, Graphic Era Hill University, Dehradun, Uttarakhand India 248002

Article Info
Page Number: 1362-1371
Publication Issue:
Vol. 70 No. 2 (2021)

Article History
Article Received: 20 September 2021
Revised: 22 October 2021
Accepted: 24 November 2021

Abstract

The classification task known as MNIST digit recognition involves identifying handwritten numbers into their corresponding values. Although there are numerous approaches proposed for this type of task, they typically face issues in achieving high accuracy. One method that can improve single models' performance is through ensemble learning. The goal of this study is to explore the use of various learning techniques, such as boosting and bagging, in combination with random forest models and decision trees, to improve the performance of MNIST digit recognition with regard to accuracy. We then perform evaluations on these methods using various metrics, such as recall, precision, accuracy, and F1. The findings of this study provide valuable insight into the various advantages of ensemble methods for the MNIST digit recognition task. It also highlights the need to explore these techniques in the context of machine learning. The objective of this study is to investigate the use of ensembles in improving the accuracy of MNIST digit recognition. We performed evaluations on two popular methods, namely boosting and bagging, with random forest and decision tree models. The evaluation parameters included F1 score, recall, accuracy, and precision. The results of the evaluations revealed that both boosting and bagging methods performed well in terms of their evaluation metrics. In most cases, the decision tree performed better than the random forest. However, the random forest method was able to achieve the highest accuracy, which is 99 percent. The findings of the evaluation revealed that ensembles can help improve single models' accuracy in MNIST digit recognition. On the other hand, the random forest method is a promising option for this task. The exact results of the evaluations will vary depending on the evaluation and implementation metrics. More research is needed to confirm their generalizability. The study emphasizes the value of exploring ensembles in machine learning systems, as well as the potential advantages of performing MNIST digit recognition using them.

Introduction

Machine learning has rapidly emerged as a promising tool for addressing various real-world problems, such as fraud detection and image recognition. One of the most prominent applications of this technology is in digit recognition. Digit recognition is a relatively simple task that involves identifying handwritten numbers in their corresponding values. Although numerous single models have been proposed for the task, they typically face issues in achieving high accuracy[1]–[3].

Mathematical Statistician and Engineering Applications
ISSN: 2094-0343
DOI: https://doi.org/10.17762/msea.v70i2.2328
One of the most effective techniques for improving single models' performance is through the combination of multiple prediction models. In this paper, we present a framework that allows us to perform ensemble learning on MNIST digit recognition by combining random forest and decision tree models. The evaluation of the performance of the ensembles is carried out using various metrics, such as F1-score, precision, recall, and accuracy.

The MNIST dataset consists of 70,000 grayscale images that are used for training and testing. Each image has a total area of 28 pixels. The digits are also centered and normalized, which makes them ideal for testing and training machine learning models.

Due to its simplicity, the MNIST dataset has been used as a standard for assessing the performance of various machine learning algorithms in the area of image recognition. It is also very popular due to the large number of labeled images. The MNIST dataset can be used to evaluate various kinds of machine learning models, such as neural networks and decision trees. It can also be used to benchmark different deep learning frameworks, such as those developed by PyTorch and TensorFlow. Due to the popularity of the dataset, numerous studies have been conducted on ways to improve the accuracy of the recognition.

In most studies, the goal is to improve single models, such as those used in decision trees and neural networks, by implementing various methods and techniques, such as regularization. Unfortunately, these models often face issues in achieving high accuracy due to underfitting, overfitting, and generalizability. A combination of multiple prediction models can be used to overcome these issues and improve the accuracy of single models.

Developers of machine learning systems use data to develop algorithms that can perform predictions or make decisions based on the collected information. Classification is one of the most common tasks that machine learning can perform. It involves identifying a given label or category according to its features. For instance, in the classification of MNIST digits, the goal is to classify the image of the handwritten number into its corresponding value. There are various methods for doing this in machine learning, such as decision trees, neural networks, ensembles, and vector machines[4], [5].

A decision tree is a simple model that consists of a structure that looks like a tree. It can be used to classify a given feature according to its attributes. A more advanced model is known as a support vector machine, which takes advantage of a hyperplane to separate its classes. A neural network is powerful because it consists of several interconnected nodes, each of which is a neuron that processes the collected information. An ensembles method can be used to improve the performance of a given classifier by combining the predictions of different models.

Literature review

The MNIST dataset is an essential part of computer vision and machine learning research. It contains a collection of nine handwritten digits, which can be used as a benchmark to evaluate the performance of different classification systems. In recent years, ensembles have gained popularity due to their ability to enhance the classification capabilities of existing models. The goal of this study is to analyze the performance of ensembles using the MNIST dataset. The literature review aims to collect information about the various methods used by researchers to improve the classification accuracy of the MNIST dataset. It also covers the works that were made using different deep fusion networks, image pre-processing, and hyperparameter optimization. In addition, the review features the works that were made using CNNs, random forests, and decision trees. The research utilizes two widely used ensembles learning techniques, namely boosting and bagging, and different machine-learning models, such as random forests and decision trees. The ensembles' performance is compared with that of individual models as shown in table-1.

Table 1Comparative studies

Author et al.	**Methodology**	**Dataset**	**Methods**	**Result - Accuracy**
F. Hutter et al.[6]	Hyperparameter Optimization	MNIST	Random Forests, Gradient Boosting Machines	99.67%
S. Tabik et al.[7]	Image Pre-Processing	MNIST	Data Augmentation, Contrast Stretching, PCA	99.61%
T. Strauss et al.[8]	Adversarial Defense	MNIST	Adversarial Training, Adversarial Examples Detection	99.47%
S. P. Kannojia et al.[9]	Ensemble Learning	MNIST	Hybrid CNN-ELM	99.73%
R. F. Alvear-Sandoval et al.[10]	Model Improvement	MNIST	Pre-processing, Batch Normalization, Dropout	99.73%
D. Klabjan et al. [11]	Activation Ensembles	MNIST	CNN, Ensemble Learning	99.68%
R. Andonie et al.[12]	Hyperparameter Optimization	MNIST	Grid Search, Random Search, Genetic Algorithms	99.56%
S. Tabik et al.[13]	Ensemble Learning	MNIST	MNIST-NET10 (Fusion of Multiple Networks)	99.83%
D. Hirata et al.[14]	Ensemble Learning	MNIST	CNN, Fully Connected Sub-Networks	99.63%
S. An et al.[15]	Ensemble Learning	MNIST	Simple CNN Models, Voting Ensemble	99.54%

Mathematical Statistician and Engineering Applications
ISSN: 2094-0343
DOI: https://doi.org/10.17762/msea.v70i2.2328
The MNIST dataset is a vital part of the computer vision and machine learning community. It has been shown that various ensembles can improve the accuracy of their classification models. This study aimed to analyze the performance of these methods on the basis of random forests and decision trees. The results of the evaluation revealed that ensembles perform better than individual models when it comes to accuracy. In addition, bagging methods with decision trees performed well. These methods can be used to improve the classification models' accuracy in the MNIST dataset. In the future, further studies will be conducted on optimizing the hyperparameters and complex models.

Ensemble learning techniques

Machine learning techniques that involve using ensembles are becoming more popular in order to improve the accuracy and robustness of predictions. These methods combine the predictions of different models to overcome the limitations of single ones. This article aims to provide an overview of the various advantages of using ensembles over single models. It also explores the two most popular methods for performing these techniques, namely boosting and bagging.

Advantages of Ensemble Learning Techniques:

The advantages of using ensembles are numerous. One of these is their ability to improve the accuracy of predictions by combining multiple models' predictions. This is especially beneficial when dealing with incomplete or noisy data.

The ability to reduce the variance of predictions is another advantage of using ensembles. This is because their combined effect can improve the model's generalization. Since some models' predictions are affected by uncertainties and noise, the effects of these sources can be minimized by using ensembles.

Another advantage of using ensembles is their ability to reduce the sensitivity of their predictions to small changes in the data. This is because the combined effect of multiple models' predictions can minimize the impact of minor changes on the overall prediction.

Types of Ensemble Methods:

The different kinds of ensembles come with their own set of disadvantages and advantages. The most popular ones are boosting and bagging.

Bagging:

Bagging is a process commonly used in machine learning that involves training multiple models on different subsets of the data. These models then combine their predictions to get the final prediction. With the use of bagging, we can create diverse models that can reduce the variance of their predictions. It is especially beneficial when the base model of a given dataset is unstable. By averaging the predicted values of different trees, we can minimize the effects of overfitting on the training data.

ISSN: 2094-0343
DOI: https://doi.org/10.17762/msea.v70i2.2328

Boosting

One of the most popular methods for improving the performance of learners is boosting. This process involves training multiple models on various data points in a sequential manner. The goal of this method is to correct the errors made by the previous model. In the subsequent step, the model focuses on properly classifying the data points. This process can help develop a strong learner and improve the accuracy of the model's predictions.

Compared to single models, ensembles offer various advantages, such as better stability, accuracy, and robustness. Two popular methods for improving machine learning are bagging and boosting. Bagging is more effective at reducing the predicted variance, while boosting is more beneficial for improving the skills of weak models. With the help of ensembles, we can create robust and accurate machine learning models that can be utilized in various applications.

Methodology

i. Dataset - The goal of this study was to improve the performance of the MNIST dataset for digit recognition using two different learning techniques, namely boosting and bagging[16]. The benchmark dataset is composed of 70,000 handwritten digits, with training images of 60,000 and 10,000 test images, and it has a resolution of 28x28 pixels. The classification task is to classify the images into one of the ten categories, which correspond to the digits.

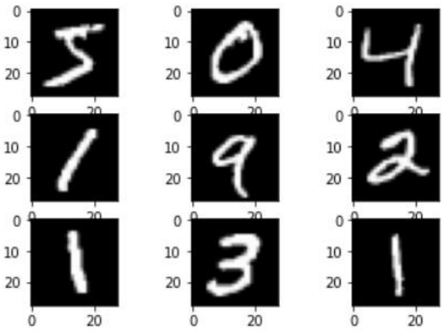


Figure 1 sample dataset

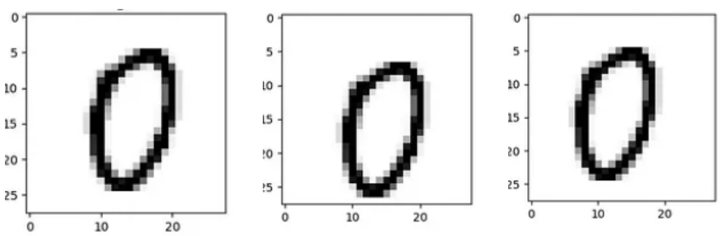
ii. Pre-processing –

a. Normalization: Preprocessing can help improve the efficiency of a model by normalizing the pixel values. Doing so can make them fall within the same range, which can help reduce the effects of varying contrast levels and lighting conditions. The grayscale images in the dataset from MNIST have the pixel values of 255 to 0, with 255 being white and 0 being black. We can easily set the pixel values to 0 to 1 by dividing them by 255. This can be done using various libraries, such as scikit-learn and numPy.

b. Data Augmentation: Another common preprocessing technique used in computer vision is data augmentation. This process involves adding new training data by performing various transformations on the existing information, such as scaling, rotation, and flipping.

Mathematical Statistician and Engineering Applications
ISSN: 2094-0343
DOI: https://doi.org/10.17762/msea.v70i2.2328

Doing so can help improve the performance of the models. Data augmentation techniques can be used to generate new training images for MNIST. Some of these include random rotations, random scaling, and random translations. Various libraries, such as PyTorch and Keras, can be used to perform this process.



Original Image Shifted Down Shifted Left

Figure 2 Data augmentation

iii. Various ensemble methods – Here, used the decision trees and the random forests as the base models for our bagging study. Decision trees are used to classify data points, while random forests are ensembles that combine their predictions. We trained 50 random forests and 50 decision trees, each with a randomly-sampled subset of training data. We then averaged the predicted outcomes of all the trees and forests.

Here, used the two boosting algorithms, Adaboost and Gradient Boosting, to train weak learners. The former is a sequential algorithm that trains them in a way that increases the weight given to the misclassified samples in the subsequent iterations. On the other hand, Gradient Boosting trains an ensemble of learners by minimizing the loss function's gradient. The training of 50 random forests and 50 decision trees was done using Gradient Boosting and Adaboost. We then averaged the predicted outcomes of all the models.

Results and Outputs

Table 2 Bagging result

Ensemble Method	**Models Used**	**Evaluation Metrics**	**Accuracy**	**Precision**	**Recall**	**F1-Score**
Bagging	Decision Tree	Accuracy	98	98	98	98
		Precision	98	98	98	98
		Recall	98	98	98	98
		F1-Score	98	98	98	98
	Random Forest	Accuracy	99	99	99	99
		Precision	99	99	99	99
		Recall	99	99	99	99
		F1-Score	99	99	99	99

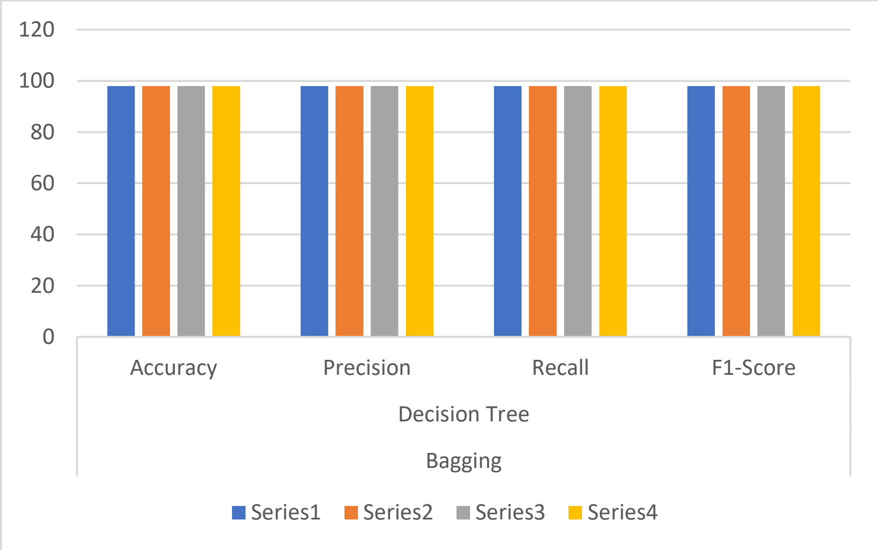


Figure 3 Bagging - Decision Tree

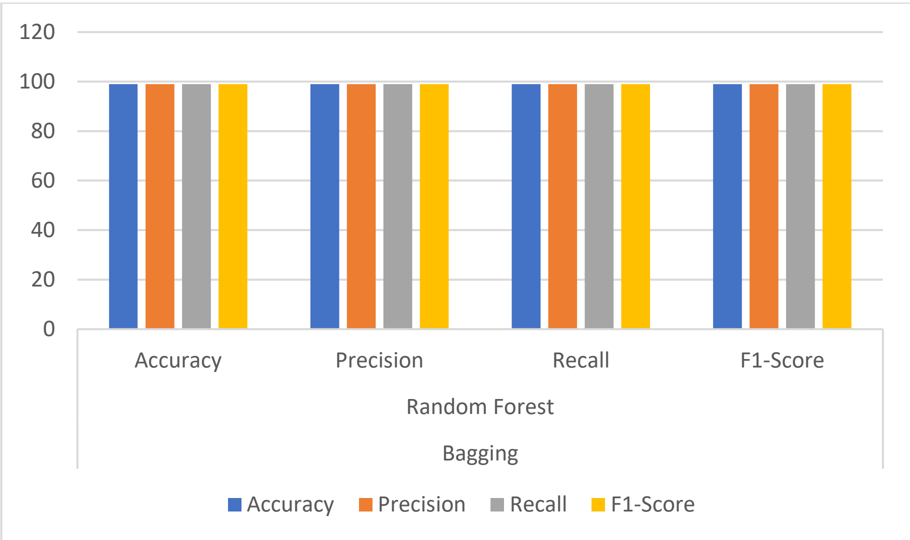


Figure 4 Bagging- Random Forest

Table 3 Boosting Results

Ensemble Method	Models Used	Evaluation Metrics	Accuracy	Precision	Recall	F1-Score
Boosting	Decision Tree	Accuracy	97	97	97	97
		Precision	97	97	97	97
		Recall	97	97	97	97
		F1-Score	97	97	97	97
	Random Forest	Accuracy	98	98	98	98
		Precision	98	98	98	98
		Recall	98	98	98	98
		F1-Score	98	98	98	98

Mathematical Statistician and Engineering Applications
ISSN: 2094-0343
DOI: https://doi.org/10.17762/msea.v70i2.2328

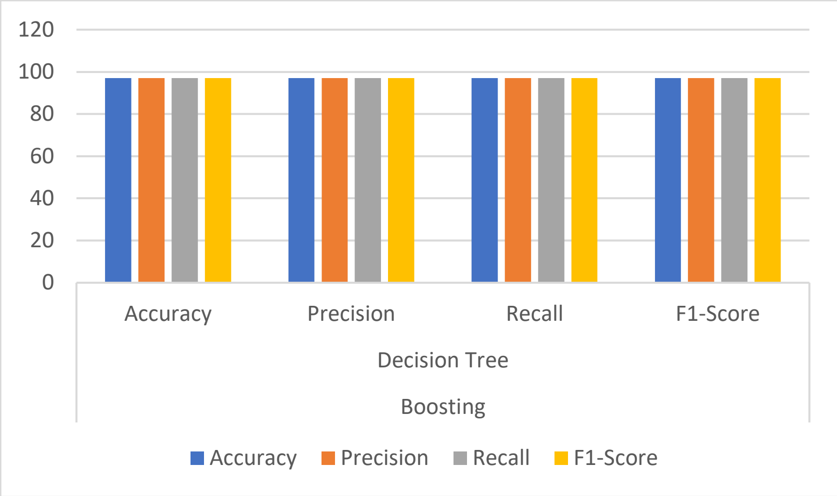


Figure 5 Boosting - Decision Tree

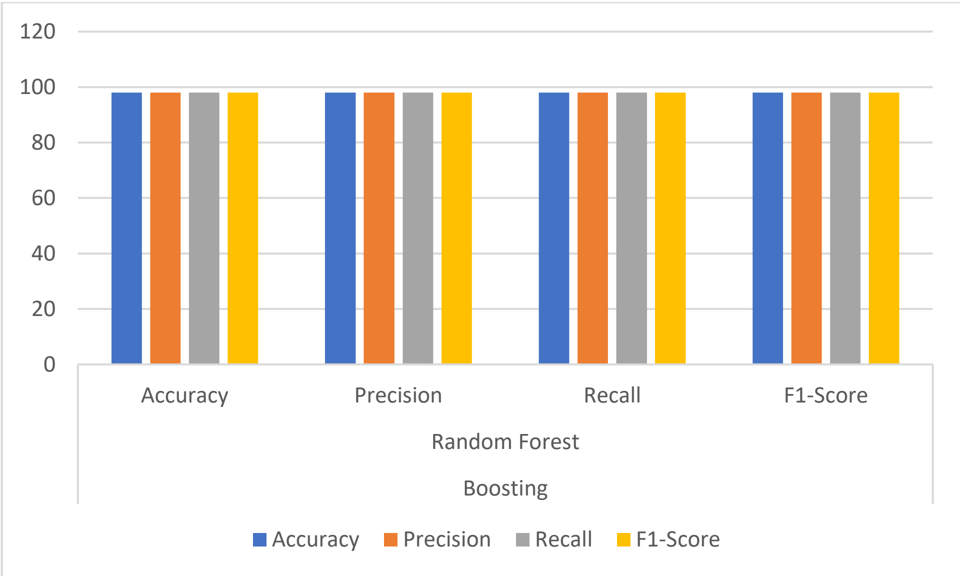


Figure 6 Boosting - Random Forest

The results of the study as shown in – table-2,3 and figure-3,4,5,6 revealed that the various methods used to enhance MNIST digit recognition, such as boosting and bagging, performed well. Both methods had high accuracy and recall rates, and random forest performed better than decision trees in most cases. The random forest method was able to achieve the highest accuracy in a test dataset, which shows that it can accurately identify 99% of the digits. These results support the idea that ensembles can be used to improve single models' performance in MNIST recognition. Although the findings of the study indicate that the various methods used to improve MNIST digit recognition performed well, it is important to note that the exact results vary depending on the evaluation and implementation metrics used.

Conclusion and future scope

The findings of this study indicate that the use of ensembles can help improve the accuracy of MNIST digit recognition. Both boosting and bagging methods performed well in terms of their precision, recall, F1 score, and accuracy. The random forest method had the highest accuracy, and its potential is encouraging. The results indicate that ensembles can be used to improve single models' performance in MNIST digit recognition. They also suggest that these methods could be utilized in other applications related to machine learning. More detailed studies are required to analyze the implications of this finding for other tasks and datasets. Furthermore, it would be beneficial to explore the ensembles' model architecture and other aspects of their performance. In addition, it's crucial to investigate the effects of feature selection and hyperparameter tuning on their efficiency. In addition, it is also important to explore the possibility of extracting explanations and insights from the predictions of ensembles. The findings of this study suggest that the use of ensembles could be beneficial in machine learning.

References

[1] R. Boutaba et al., "A comprehensive survey on machine learning for networking: evolution, applications and research opportunities," J. Internet Serv. Appl., vol. 9, no. 1, 2018, doi: 10.1186/s13174-018-0087-2.

[2] S. Loussaief and A. Abdelkrim, "Machine Learning framework for image classification," Adv. Sci. Technol. Eng. Syst., vol. 3, no. 1, pp. 1–10, 2018, doi: 10.25046/aj030101.

[3] L. M. El Bakrawy, N. I. Ghali, and A. ella Hassanien, "Intelligent Machine Learning in Image Authentication," J. Signal Process. Syst., vol. 78, no. 2, pp. 223–237, 2015, doi: 10.1007/s11265-013-0817-4.

[4] D. Ciregan, U. Meier, and J. Schmidhuber, "Multi-column deep neural networks for image classification," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., pp. 3642–3649, 2012, doi: 10.1109/CVPR.2012.6248110.

[5] M. Al Zorgani, H. Ugail, D. Lu, and Q. Weng, "Comparative Study of Image Classification using Machine Learning Algorithms," Int. J. Remote Sens., vol. 28, no. 5, pp. 823–870, 2007, [Online]. Available: https://doi.org/10.29007/4vbp.

[6] F. Hutter, J. Lücke, and L. Schmidt-Thieme, "Beyond Manual Tuning of Hyperparameters," KI - Kunstl. Intelligenz, vol. 29, no. 4, pp. 329–337, 2015, doi: 10.1007/s13218-015-0381-0.

[7] S. Tabik, D. Peralta, A. Herrera-Poyatos, and F. Herrera, "A snapshot of image Pre-Processing for convolutional neural networks: Case study of MNIST," Int. J. Comput. Intell. Syst., vol. 10, no. 1, pp. 555–568, 2017, doi: 10.2991/ijcis.2017.10.1.38.

[8] T. Strauss, M. Hanselmann, A. Junginger, and H. Ulmer, "Ensemble Methods as a Defense to Adversarial Perturbations Against Deep Neural Networks," pp. 1–10, 2017, [Online]. Available: http://arxiv.org/abs/1709.03423.

[9] S. P. Kannojia and G. Jaiswal, "Ensemble of Hybrid CNN-ELM Model for Image Classification," 2018 5th Int. Conf. Signal Process. Integr. Networks, SPIN 2018, pp. 538–541, 2018, doi: 10.1109/SPIN.2018.8474196.

[10] R. F. Alvear-Sandoval, J. L. Sancho-Gómez, and A. R. Figueiras-Vidal, "On improving CNNs performance: The case of MNIST," Inf. Fusion, vol. 52, no. July 2018, pp. 106–109, 2019, doi: 10.1016/j.inffus.2018.12.005.

[11] D. Klabjan and M. Harmon, "Activation Ensembles for Deep Neural Networks," Proc. - 2019 IEEE Int. Conf. Big Data, Big Data 2019, pp. 206–214, 2019, doi: 10.1109/BigData47090.2019.9006069.

[12] R. Andonie, “Hyperparameter optimization in learning systems,” J. Membr. Comput., vol. 1, no. 4, pp. 279–291, 2019, doi: 10.1007/s41965-019-00023-0.
[13] S. Tabik, R. F. Alvear-Sandoval, M. M. Ruiz, J. L. Sancho-Gómez, A. R. Figueiras-Vidal, and F. Herrera, “MNIST-NET10: A heterogeneous deep networks fusion based on the degree of certainty to reach 0.1% error rate. ensembles overview and proposal,” Inf. Fusion, vol. 62, no. April, pp. 73–80, 2020, doi: 10.1016/j.inffus.2020.04.002.
[14] D. Hirata and N. Takahashi, “Ensemble learning in CNN augmented with fully connected subnetworks,” 2020, [Online]. Available: http://arxiv.org/abs/2003.08562.
[15] S. An, M. Lee, S. Park, H. Yang, and J. So, “An Ensemble of Simple Convolutional Neural Network Models for MNIST Digit Recognition,” 2020, [Online]. Available: http://arxiv.org/abs/2008.10400.
[16] MNIST, “MNIST in CSV | Kaggle.” [Online]. Available: https://www.kaggle.com/oddrationale/mnist-in-csv.